Available online at www.sciencedirect.com
SCIENCE DIRECT®

J. Math. Anal. Appl. 301 (2005) 336–353

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

A new weighted metric: the relative metric II

Peter A. Hästö [1]

Department of Mathematics, Norwegian University of Science and Technology, 7491 Trondheim, Norway

Received 27 May 2003
Available online 23 September 2004
Submitted by L. Chen

Abstract

In the first part of this investigation we generalized a weighted distance function of R.-C. Li's and found necessary and sufficient conditions for it being a metric. In this paper some properties of this so-called M-relative metric are established. Specifically, isometries and quasiconvexity results are derived. We also illustrate connections between our approach and generalizations of the hyperbolic metric.

Keywords: Relative metric; Weighted metric; Hyperbolic metric; Quasihyperbolic metric

1. Preliminaries and main results

The p-relative metric, for $p \in [1, \infty]$, was introduced by R.-C. Li [6] in connection with studying perturbation problems of matrices. Li considered the metric only on the real line, but later A. Berglund [1] proved that it is a metric also in the complex plane. The author proved in the first part of this investigation, [4], that we can think of these metrics in a more general setting by replacing p with a symmetric function M. In this paper, some properties of these so-called M-relative metrics are established. Some further M-relative metrics were established in [5].

E-mail address: peter.hasto@helsinki.fi.
[1] Supported in part by the Academy of Finland and the Finnish Academy of Science and Letters (Viljo, Yrjö and Kalle Väisälä's Fund).

doi:10.1016/j.jmaa.2004.07.034

In this section we introduce the M-relative metric and state the main results. In order to do this, we have to introduce some notation—for a fuller account the reader should consult [4, Section 2].

We denote $\mathbb{R}^+ = [0, \infty)$. An increasing function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is said to be *moderately increasing* if $f(t)/t$ is decreasing on $(0, \infty)$. A function $P : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ of two variables is moderately increasing if both $P(x, \cdot)$ and $P(\cdot, x)$ are moderately increasing for each fixed $x \in \mathbb{R}^+$.

If $P : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ satisfies

$$\max\{x^\alpha, y^\alpha\} \geqslant P(x, y) \geqslant \min\{x^\alpha, y^\alpha\}$$

for all $x, y \in \mathbb{R}^+$, $\alpha > 0$, then it is called an *α-quasimean*. A 1-quasimean is called a *mean*. We will need the following family of quasimeans, related to the Stolarsky mean [8]:

$$S_p(x, y) := (1 - p)\frac{x - y}{x^{1-p} - y^{1-p}}, \qquad S_p(x, x) = x^p, \qquad 0 < p < 1,$$

$$S_1(x, y) := L(x, y) := \frac{x - y}{\log x - \log y}, \qquad S_1(x, x) = x.$$

Throughout this paper we will denote by M a symmetric function, $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ with $M(x, y) > 0$ if $xy > 0$, $x, y \in \mathbb{R}^+$. In the special case $M(x, y) = f(x)f(y)$ this means that we assume $f(x) > 0$ when $x > 0$. By the *M-relative distance* (in $\mathbb{R}^n$) we mean the function

$$\rho_M(x, y) := \frac{|x - y|}{M(|x|, |y|)},$$

where $x, y \in \mathbb{R}^n$ (here we define $0/0 = 0$). We will use the convention $M(x, y) := M(|x|, |y|)$ (and $f(x) := f(|x|)$, when $M(x, y) = f(x)f(y)$). If ρ_M is a metric, it is called the *M-relative metric*. If $M(x, y) = 0$ for some $x, y \in [0, \infty)$, we will use the convention that "ρ_M is a metric in $\mathbb{R}^n$" means that ρ_M is a metric in $\mathbb{R}^n \setminus \{0\}$. The main results of the first part of this investigation are summarized in the next theorem. (The results in that paper were actually derived in the more general setting of Ptolemaic normed spaces. However, for simplicity we restrict our attention to $\mathbb{R}^n$ in this paper.)

Theorem 1.1 [4, Sections 1 and 3]. *Let $n \geqslant 2$ and consider the space $\mathbb{R}^n$.*

(1) *Assume that M is moderately increasing. Then ρ_M is a metric in $\mathbb{R}^n$ if and only if it is a metric in $\mathbb{R}$.*
(2) *Let M be an α-quasimean. Then ρ_M is a metric in $\mathbb{R}$ if $M(x, 1)/S_\alpha(x, 1)$ is increasing in x for $x \geqslant 1$. If ρ_M is a metric in $\mathbb{R}$ then $M(x, 1) \geqslant S_\alpha(x, 1)$ for $x \geqslant 1$.*
(3) *Assume that $M(x, y) = (x^p + y^p)^{q/p}$ for $p, q > 0$. Then ρ_M is denoted by $\rho_{p,q}$, and it is a metric in $\mathbb{R}^n$ if and only if $0 < q \leqslant 1$ and $p \geqslant \max\{1 - q, (2 - q)/3\}$.*
(4) *Let $M(x, y) = f(x)f(y)$ where $f : \mathbb{R}^+ \to (0, \infty)$. Then ρ_M is a metric in $\mathbb{R}^n$ if and only if f is moderately increasing and convex.*

In this paper we establish some properties of relative metrics that help us get an intuitive grasp of what the metrics look like. We first consider bi-Lipschitz mappings. The idea with such studies is that if (all) bi-Lipschitz mappings distort a metric to a similar degree, then the metric is somehow similar to the Euclidean metric. Like the first part of the investigation, this paper is organized along three threads—one general and two special ones. In the general case, we show in Section 2 that Euclidean bi-Lipschitz mappings are bi-Lipschitz in the relative metric, provided the mapping fixes the origin and M is moderately increasing.

A second regularity property that we will consider is quasiconvexity, which plays an important role for instance in J. Väisälä's study of quasiconformality in metric spaces [10]. Quasiconvexity measures how far from equality we are in the triangle inequality, and is thus more of an intrinsic measure of regularity than bi-Lipschitz mappings. We define a metric ρ_M (actually a metric space, $(\mathbb{R}^n, \rho_M)$) to be c-quasiconvex if there exists a path γ joining x and y such that $\ell_M(\gamma) \leqslant c\rho_M(x, y)$, where $\ell_M(\gamma)$ denotes the length of γ in the metric ρ_M. For instance, if $G \subset \mathbb{R}^n$ is convex then $(G, |\cdot|)$ is 1-quasiconvex, whereas $(D, |\cdot|)$ is not quasiconvex in $\mathbb{R}^2$ for $D := B^2 \setminus [0, 1)$.

In Section 4 we show how one can calculate the quasiconvexity constant of a relative metric when M is a quasimean. In the special cases we can prove a bit more:

Theorem 1.2. *Let $\rho_{p,q}$ denote the (p, q)-relative metric as in Theorem 1.1(3) and assume that $n \geqslant 2$. Then the (p, q)-relative metric is quasiconvex in $\mathbb{R}^n$ if and only if $q < 1$, in which case it is $c_{p,q}$-quasiconvex, where*

$$\frac{2^{-q/p}}{1-q} \leqslant c_{p,q} \leqslant \frac{\max\{2^{q(1-1/p)}, 1\}}{1-q}.$$

Theorem 1.3. *Let $M(x, y) = f(x)f(y)$ and assume that ρ_M is a metric in $\mathbb{R}^n$. If $n \geqslant 2$, ρ_M is c-quasiconvex for some $c \leqslant \sqrt{\pi^2/4 + 4}$.*

This paper also contains an explicit formula for the α-quasihyperbolic metric, k_α, in the domain $\mathbb{R}^n \setminus \{0\}$ which might be of independent interest (the α-quasihyperbolic metric is defined in the beginning of Section 4).

Theorem 1.4. *For $n \geqslant 2$ and $0 < \alpha < 1$, we have*

$$k_\alpha(x, y) = \frac{1}{\beta}\sqrt{|x|^{2\beta} + |y|^{2\beta} - 2|x|^\beta |y|^\beta \cos\beta\theta},$$

where $\alpha + \beta = 1$ and θ is the angle $\widehat{x0y}$. In particular, as $\alpha \to 1$,

$$k_\alpha(x, y) \to \sqrt{\theta^2 + \log^2(|x|/|y|)},$$

the well-known expression for the quasihyperbolic metric in $\mathbb{R}^n \setminus \{0\}$ [11, 3.11].

In the last section, we consider how the relative-metric approach may be applied to extending the hyperbolic metric in $\mathbb{R}^n$ for $n \geqslant 3$. The hyperbolic metric is a basic tool in analysis, and so its study needs lite further motivation. We consider a generalization of the hyperbolic metric proposed by M. Vuorinen in [11, 3.25, 3.26]. We conclude that this metric cannot be interpreted in the present framework of relative metrics and finally use a different method to prove that ρ_M satisfies the triangle inequality.

2. Bi-Lipschitz mappings and ρ_M

Proposition 2.1. *Let M be moderately increasing and ρ_M be a metric in $\mathbb{R}^n$. If $g:\mathbb{R}^n \to \mathbb{R}^n$ is L-bi-Lipschitz, $L \geqslant 1$, with respect to the norm $|\cdot|$ and $g(0)=0$, then g is L^3-bi-Lipschitz with respect to the metric ρ_M.*

Proof. Assume first that $x, y \neq 0$. Since M is moderately increasing

$$\rho_M\big(g(x), g(y)\big) = \frac{|g(x)-g(y)|}{M(g(x), g(y))} \leqslant \frac{L|x-y|}{M(x/L, y/L)} \leqslant L^3 \frac{|x-y|}{M(x,y)},$$

where the last inequality follows since

$$\frac{M(x/L, y/L)}{xy/L^2} \geqslant \frac{M(x, y/L)}{xy/L} \geqslant \frac{M(x,y)}{xy},$$

by the moderation condition. On the other hand, if $y=0$ and $M(g(x),0) > 0$ then

$$\rho_M\big(g(x), 0\big) = \frac{|g(x)|}{M(g(x),0)} \leqslant \frac{L|x|}{M(x/L, 0)} \leqslant L^2 \frac{|x|}{M(x,0)}.$$

In the case $M(g(x),0)=0$ the point 0 need not be considered in the triangle inequality by convention. The lower Lipschitz bound follows similarly. □

Remark 2.1. It is clear that the condition $g(0)=0$ in Lemma 2.1 is essential. For the translation $x \mapsto x+a$ is 1-bi-Lipschitz in the norm $|\cdot|$. But if $M(x,y)=x+y$ then

$$\lim_{\epsilon\to 0} \frac{\rho_M(-\epsilon, \epsilon)}{\rho_M(a-\epsilon, a+\epsilon)} = \infty,$$

in $\mathbb{R}$, hence the translation is not bi-Lipschitz in ρ_M. Note also that the condition $g(0)=0$ can be understood in terms of the generalization of the relative metric presented in [4, Section 6]: the ρ_M is finite in $\mathbb{R}^n \setminus \{0\}$ if M is moderately increasing (and $M \not\equiv 0$) and hence the relevant class of mappings are from $\mathbb{R}^n \setminus \{0\}$ to $\mathbb{R}^n \setminus \{0\}$, which, when continuously extended to $\mathbb{R}^n$, have $g(0)=0$.

The next lemma shows that ρ_M bi-Lipschitz mappings are in general quasiconformal. For background information on quasiconformal mappings the reader may consult [9]. This material will not be needed in any subsequent results, however.

Proposition 2.2. *Let M be moderately increasing with $M \not\equiv 0$. If ρ_M is a metric in $\mathbb{R}^n$ and $g:\mathbb{R}^n \to \mathbb{R}^n$ is L-bi-Lipschitz, $L \geqslant 1$, with respect to the metric ρ_M then g is quasiconformal in $\mathbb{R}^n$ with linear dilatation coefficient less than or equal to L^2.*

Proof. We will first prove that g is continuous in $\mathbb{R}^n \setminus \{0, g^{-1}(0)\}$ with respect to $|\cdot|$. Recall that by the definition of M we have $M(x,y) > 0$ unless $|x||y| = 0$. Since M is moderately increasing it is continuous in $(0,\infty)\times(0,\infty)$ by [4, Lemma 2.4].

Fix a point $x \in \mathbb{R}^n$, $x \neq 0$, such that $g(x) \neq 0$. Since ρ_M is a metric and g is bi-Lipschitz with respect to ρ_M it follows that g is injective. Hence there exists a closed set U_x such that $x \in \operatorname{int} U_x$ (the interior of U_x) and $0 \notin U_x$. We will derive an upper bound for $|g(y)|$ when y in a neighborhood of x contained in U_x. If $|g(y)| \geqslant |g(x)|$ then the inequality

$$\big|g(y)\big| - \big|g(x)\big| \leqslant \big|g(x) - g(y)\big| \leqslant L \frac{|g(y)|}{|g(x)|} M\big(g(x), g(x)\big) \rho_M(x, y), \tag{1}$$

implies that

$$\big|g(y)\big| \leqslant \big|g(x)\big| \left(1 - L \frac{M(g(x), g(x))}{|g(x)|} \rho_M(x, y)\right)^{-1}$$

for y such that $\rho_M(x, y) < |g(x)|/(LM(g(x), g(x)))$ (this inequality holds in a neighborhood of x since M is continuous). If $|g(y)| \leqslant |g(x)|$ then the previous estimate holds trivially. Hence it follows from 1 that

$$\big|g(x) - g(y)\big| \leqslant \frac{L'}{1 - L'\rho_M(x, y)/|g(x)|} \rho_M(x, y),$$

where $L' := LM(g(x), g(x))$. For fixed x, it is clear that L' is bounded. From this it follows that $g(y) \to g(x)$ as $y \to x$, since $\rho_M(x, y) \to 0$. Hence g is continuous in $\mathbb{R}^n \setminus \{0, g^{-1}(0)\}$.

Let $x \notin \{0, g^{-1}(0)\}$, $y, z \in U_x$ and $|x - z| = |y - x| = r$. Then

$$\frac{|g(x) - g(z)|}{|g(y) - g(x)|} \leqslant L^2 \frac{M(g(x), g(z)) M(y, x)}{M(g(y), g(x)) M(x, z)}.$$

By the continuity of M and g the right-hand side tends to L^2 as $r \to 0$. Hence we have proved that g is quasiconformal in $\mathbb{R}^n \setminus \{0, g^{-1}(0)\}$. But then g is quasiconformal in $\mathbb{R}^n$ by well-known continuation results (see, e.g., [9, Theorem 35.1]). $\square$

Remark 2.2. If M and g are as in the previous lemma and additionally $M(x, 0) = 0$ for every $x > 0$ then $g(0) = 0$. For the bi-Lipschitz condition

$$\frac{1}{L} \frac{|x - y|}{M(x, y)} \leqslant \frac{|g(x) - g(y)|}{M(g(x), g(y))} \leqslant L \frac{|x - y|}{M(x, y)}$$

implies that $M(x, y)$ and $M(g(x), g(y))$ are simultaneously 0. Therefore

$$\begin{aligned} |x||y| = 0 \quad &\Leftrightarrow \quad M(x, y) = 0 \quad \Leftrightarrow \quad M\big(g(x), g(y)\big) = 0 \\ &\Leftrightarrow \quad \big|g(x)\big|\big|g(y)\big| = 0, \end{aligned}$$

which implies $g(0) = 0$.

Corollary 2.3. *If M is moderately increasing with $M \not\equiv 0$ and $g : \mathbb{R}^n \to \mathbb{R}^n$ is a ρ_M-isometry then g is conformal.*

Remark 2.4. The mapping $g(x) = |x|x$ is 2-bi-Lipschitz in the $\rho_{\infty,1}$ metric (which is ρ_M with $M(x, y) = \max\{x, y\}$) but is not Lipschitz with respect to the Euclidean metric (ρ_M with $M \equiv 1$). The spherical chordal metric q is a relative metric with $M(x, y) = \sqrt{1 + x^2}\sqrt{1 + y^2}$ and the inversion $x \mapsto x/|x|^2$ is a q-isometry. However, this inversion is certainly not Lipschitz with respect to the Euclidean metric. These examples show that the class of ρ_M-Lipschitz mappings depends on M in a non-trivial way.

3. α-quasihyperbolic metrics

The length of a (rectifiable) path $\gamma:[0,l]\to\mathbb{R}^n$ in the metric ρ_M with continuous M is defined by

$$\ell_M(\gamma) := \lim_{n\to\infty}\sum_{i=0}^{n}\rho_M\big(\gamma(t_i),\gamma(t_{i+1})\big),$$

where $t_i < t_{i+1}$, $t_0 = 0$, $t_n = l$ and $\max\rho_M(\gamma(t_{i+1}),\gamma(t_i))\to 0$. If γ is any path connecting x and y in $\mathbb{R}^n$ then $\rho_M(x,y)\leqslant\ell_M(\gamma)$ by the triangle inequality.

Let M be an α-quasimean ($0<\alpha\leqslant 1$). By taking the infimum over all rectifiable paths joining x and y, we conclude that

$$\rho_M(x,y)\leqslant\inf_\gamma \ell_M(\gamma) = \inf_\gamma\int\frac{ds}{|\gamma(s)|^\alpha} =: k_\alpha(x,y), \tag{2}$$

since $M(x,x+\epsilon)\geqslant x^\alpha$ for $\epsilon>0$. Here k_α stands for the α-quasihyperbolic metric, which was introduced in [3]. More precisely, it is the α-quasihyperbolic metric in the domain $G=\mathbb{R}^n\setminus\{0\}$. In this section we will derive an explicit expression for $k_\alpha(x,y)$, which will be used to study quasiconvexity in the next section.

Proof of Theorem 1.4. It is clearly sufficient to limit ourselves to the case $n=2$ in this proof. It is also clear that the geodesic can be parametrized by $(r(\theta),\theta)$ in polar coordinates. The kernel of the integral (2) then becomes $r^{-\alpha}\sqrt{(r')^2+r^2}$, where $r'=dr/d\theta$. Then the Euler equation (cf. [2, p. 36 (5)]) tells us that the geodesic satisfies the differential equation

$$r^{-\alpha}\sqrt{(r')^2+r^2}-\frac{r^{-\alpha}(r')^2}{\sqrt{(r')^2+r^2}} = c_1.$$

Since c_1 is independent of r, one easily sees that $c_1\neq 0$. Then the equation is equivalent to $r^\beta/c_1=\sqrt{((\log r)')^2+1}$.

To solve this equation, we change variables by substituting $y:=\log r$. The equation then becomes $e^{\beta y}=c_1\sqrt{(y')^2+1}$, where $y'=dy/d\theta$. We introduce an auxiliary parameter, t, by $\sinh t = y'$. Then $e^{\beta y}=c_1\cosh t$ and

$$d\theta = \frac{\partial y/\partial t}{\partial y/\partial\theta}\,dt = \frac{dt}{\beta\cosh t}.$$

Solving this equation gives $\tan((\beta\theta+c_2)/2)=e^t$, hence

$$r(\theta)^\beta = \frac{c_1}{2}\left(\tan\big((\beta\theta+c_2)/2\big)+\frac{1}{\tan((\beta\theta+c_2)/2)}\right) = \frac{c_1}{\sin(\beta\theta+c_2)}.$$

Let us now calculate the distance in the k_α metric between 1 and $re^{i\theta_1}$, where $r \geqslant 1$ and $0 \leqslant \theta_1 \leqslant \pi$, using the formula for the geodesic (denoted by γ):

$$k_\alpha\big(1, re^{i\theta_1}\big) = \int_\gamma \frac{\sqrt{(r')^2 + r^2}}{r^\alpha}\, d\theta = \int_0^{\theta_1} \frac{c_1}{\sin^2(\beta\theta + c_2)}\, d\theta$$
$$= \frac{c_1}{\beta}\big(\cot c_2 - \cot(\beta\theta_1 + c_2)\big).$$

It remains to express c_1 and c_2 in terms of the boundary values:

$$\sin c_2 = c_1, \qquad r^\beta \sin(\beta\theta_1 + c_2) = c_1.$$

These equations imply that

$$c_1 = \frac{r^\beta \sin\beta\theta_1}{\sqrt{1 + r^{2\beta} - 2r^\beta \cos\beta\theta_1}},$$

from which it follows that

$$k_\alpha(1, re^{i\theta_1}) = \frac{1}{\beta}\Big(\sqrt{r^{2\beta} - c_1^2} \pm \sqrt{1 - c_1^2}\Big) = \frac{r^\beta |r^\beta - \cos\beta\theta_1| \pm |r^\beta \cos\beta\theta_1 - 1|}{\beta\sqrt{1 + r^{2\beta} - 2r^\beta\cos\beta\theta_1}},$$

where $\pm$ is a plus when c_2 is greater than $\pi/2$ and a minus when it is not. This means that effectively the absolute value is disregarded and the $\pm$ sign is a minus sign since c_2 is greater than $\pi/2$ exactly when $r^\beta \cos\beta\theta_1 \geqslant 1$.

Then

$$r^\beta\big|r^\beta - \cos\beta\theta_1\big| \pm \big|r^\beta\cos\beta\theta_1 - 1\big| = r^\beta\big(r^\beta - \cos\beta\theta_1\big) - \big(r^\beta\cos\beta\theta_1 - 1\big)$$
$$= 1 + r^{2\beta} - 2r^\beta\cos\beta\theta_1$$

from which the claim follows. □

Remark 3.1. To get a picture of what k_α looks like we consider how the distance between points changes as α changes. Since k_α is β-homogeneous and spherically symmetric, we assume that $y = 1$. Consider first the case when x is a real number greater than one. Then $k_\alpha(x, 1) = (x^\beta - 1)/\beta$. This is an increasing function with respect to β. Consider now another point $z \in S^{n-1}(0, 1)$. Then $k_\alpha(z, 1) = \sqrt{2(1 - \cos\beta\theta)}/\beta$. This is decreasing in β. Hence, intuitively speaking, increasing α increases angular distance but decreases radial distance. Note that these considerations imply, in particular, that k_α is not monotone in α.

Corollary 3.1. *Let $\alpha + \beta = 1$ with $0 \leqslant \alpha < 1$. Then $k_\alpha(x, y) \leqslant (|x|^\beta + |y|^\beta)/\beta$.*

Lemma 3.2. *Let $\alpha + \beta = 1$ with $0 \leqslant \alpha < 1$. Then*

$$\frac{|x|^\beta - |y|^\beta}{\beta(|x| - |y|)} \leqslant \frac{k_\alpha(x, y)}{|x - y|} \leqslant \frac{k_\alpha(-|x|, |y|)}{|x| + |y|} \leqslant \frac{|x|^\beta + |y|^\beta}{\beta(|x| + |y|)} \leqslant \frac{2^\alpha}{\beta}|x - y|^{-\alpha}.$$

Let $t \in \mathbb{R}^+$. There is equality in the first inequality for $x = ty$, in the second for $x = -ty$, in the third for $x = -ty$ and $\beta = 1$ and in the fourth for $x = -y$.

Proof. The third and fourth inequalities are trivial and to prove the first two it suffices to show that

$$\frac{k_\alpha(re^{i\theta}, 1)}{|re^{i\theta} - 1|}$$

is increasing in θ for $r \geqslant 1$ and $0 \leqslant \theta \leqslant \pi$. Using the explicit formula for k_α from Theorem 1.4 we need to show that

$$\frac{1 + r^{2\beta} - 2r^\beta \cos\beta\theta}{\beta(1 + r^2 - 2r\cos\theta)}$$

is increasing in θ. We differentiate the equation with respect to θ and see that this follows if we show that

$$\big(1 + r^{2\beta} - 2r^\beta \cos\beta\theta\big)\big/\big(\beta r^\beta \sin\beta\theta\big)$$

is increasing in β.

When we differentiate this equation with respect to β, we see that it suffices to show that

$$(s - 1/s)\log s \sin x + 2x + \sin 2x \geqslant (s + 1/s)(x\cos x + \sin x), \tag{3}$$

where we have denoted $s := r^\beta \geqslant 1$ and $0 \leqslant x := \beta\theta \leqslant \pi$. Since the cases $x = 0$ and $x = \pi$ are clear, we will assume that $0 < x < \pi$. The inequality holds in (3) for $s = 1$ since $x - \sin x \geqslant \cos x(x - \sin x)$ for $x \geqslant 0$. Differentiating (3) with respect to s leads to

$$\big(s^2 + 1\big)\log s + s^2 - 1 \geqslant \big(s^2 - 1\big)(x\cos x/\sin x + 1).$$

Since $x/\tan x \leqslant 1$ for $0 < x < \pi$, it suffices to show that $\log s \geqslant 1 - 2/(s^2 + 1)$, which follows since $2/(s^2 + 1) + \log s$ is increasing in s. □

Remark 3.2. It would be interesting to see how the above estimates for k_α generalize to other domains than $\mathbb{R}^n \setminus \{0\}$.

4. Quasiconvexity

In this section we will assume that $n \geqslant 2$.

Theorem 4.1. *Let M be an α-quasimean, $0 < \alpha \leqslant 1$, such that ρ_M is a metric. Then ρ_M is quasiconvex if and only if*

$$c_M := \sup_{x \geqslant 0,\, y > 0} \frac{k_\alpha(xe_1, -ye_1)}{x + y} M(x, y) < \infty,$$

in which case it is c_M-quasiconvex.

Proof. The claim follows directly from the second inequality in Lemma 3.2, since $\inf_\gamma \ell_M(\gamma) = k_\alpha(z, w)$ for $z, w \in \mathbb{R}^n$, by definition. □

Corollary 4.1. *Let M be α-homogeneous, $0 < \alpha \leqslant 1$, with $M(1,1) = 1$ such that ρ_M is a metric in $\mathbb{R}^n$. Then ρ_M is quasiconvex if and only if*

$$c_M := \sup_{r \geqslant 1} \frac{k_\alpha(re_1, -e_1)}{r+1} M(r,1) < \infty,$$

in which case it is c_M-quasiconvex.

Let us define the power mean, A_p, for $p > 0$, of $x, y \in \mathbb{R}^+$ by taking $A_p(x,y) := ((x^p + y^p)/2)^{1/p}$.

Corollary 4.2. *Let M be α-homogeneous, $0 < \alpha < 1$, with $M(1,1) = 1$ such that ρ_M is a metric in $\mathbb{R}^n$. Then ρ_M is $2^\alpha/(1-\alpha)$-quasiconvex. If $M \leqslant A_p^\alpha$, $p > 0$, then ρ_M is $c_{p,\alpha}$-quasiconvex, where*

$$c_{p,\alpha} := \frac{\max\{2^{\alpha(1-1/p)}, 1\}}{1-\alpha}.$$

Proof. Let us first consider $M = A_1^\alpha$. Then, by Corollary 4.1 and Lemma 3.2,

$$c_M \leqslant \sup_{r \geqslant 1} \frac{r^{1-\alpha}+1}{(1-\alpha)(r+1)} \left(\frac{r+1}{2}\right)^\alpha = \frac{1}{2^\alpha(1-\alpha)} \sup_{r \geqslant 1} \frac{r^{1-\alpha}+1}{(r+1)^{1-\alpha}} = \frac{1}{1-\alpha},$$

since $(r^{1-\alpha}+1)(r+1)^{\alpha-1}$ is decreasing.

Since $A_p \leqslant \max\{2^{1-1/p}, 1\} A_1$, the second claim follows. Since

$$M(x,1) \leqslant \max\{1, x^\alpha\} \leqslant \{2A_1(x,1)\}^\alpha$$

for every α-homogeneous M, the first claim also follows. □

Proof of Theorem 1.2. The upper bound follows from Corollary 4.2. For the lower bound let $r \to \infty$ in Corollary 4.1. □

Corollary 4.3. *$\rho_{p,1/2}$ is $\max\{\sqrt{2}, 2^{1-1/(2p)}\}$-quasiconvex, where the constant is the best possible.*

Proof. Setting $\alpha = 1/2$ in Corollary 4.1 yields

$$c_M = \sup_{r \geqslant 1} 2^{1-1/(2p)} \sqrt{(r^p+1)^{1/p}/(r+1)},$$

from which the claim follows since $(r^p+1)^{1/p}/(r+1)$ is increasing for $p \geqslant 1$ and decreasing for $p \leqslant 1$. □

Proof of Theorem 1.3. Since $f(0) > 0$ we may assume without loss of generality that $f(0) = 1$. Let us fix the points x and y with $|x| \geqslant |y| > 0$. Denote by γ_1 the path which is radial from x to $(|y|/|x|)x$ and then circular (with radius $|y|$) about the origin to y and by γ_2 the path which is first circular (with radius $|x|$) and then radial from $(|x|/|y|)y$ to y.

In what follows, we will denote $|x|$ by x and similarly for y and z, since there is no danger of confusion. We derive estimates for the lengths of the γ_i:

$$\min\{\ell_M(\gamma_1), \ell_M(\gamma_2)\} \leqslant \theta \min\left(\frac{x}{f(x)^2}, \frac{y}{f(y)^2}\right) + \int_y^x \frac{dz}{f(z)^2},$$

where θ is the angle $\widehat{x0y}$. Since f is moderately increasing and convex we find that

$$f(z) \geqslant \max\{1 + z(f(y)-1)/y, zf(x)/x\}$$

for $z \in [y,x]$. Let $z_0 \in \mathbb{R}^+$ be such that $1 + z_0(f(y)-1)/y = z_0 f(x)/x$. Then

$$\int_y^x \frac{dz}{f(z)^2} \leqslant \int_y^{z_0} \frac{dz}{\{1+z(f(y)-1)/y\}^2} + \int_{z_0}^x \frac{dz}{\{zf(x)/x\}^2}$$
$$\leqslant \frac{2x}{f(x)} - \frac{y}{f(y)} - \frac{x}{f(x)^2}\left(\frac{x}{y}(f(y)-1)+1\right).$$

We will next show that

$$\frac{2x}{f(x)} - \frac{y}{f(y)} - \frac{x}{f(x)^2}\left(\frac{x}{y}(f(y)-1)+1\right) \leqslant \frac{2(x-y)}{f(x)f(y)}.$$

To see that this inequality holds, multiply by $f(x)^2 f(y)$ and rearrange:

$$2(x-y)f(x) - 2xf(y)f(x) + yf(x)^2 \geqslant \left(\frac{x}{y}(f(y)-1)+1\right)xf(y).$$

Notice that the right-hand side is independent of $f(x)$ whereas the left-hand side is increasing in $f(x)$ since

$$y(f(x)-1) = (y-0)(f(x)-f(0)) \geqslant (x-0)(f(y)-f(0)) = x(f(y)-1),$$

which follows from the convexity of f. The inequality then follows, when we insert the minimum value for $f(x)$, that is $x(f(y)-1)/y+1$ and use $y(f(x)-1) \geqslant x(f(y)-1)$ again.

In the case $y=0$ which was excluded above, one easily derives the estimate

$$\ell_M([0,x]) \leqslant \frac{2f(x)-1}{f(x)^2}x \leqslant \frac{2x}{f(x)} = \frac{2(x-y)}{f(x)f(y)},$$

where $[0,x]$ denotes the segment with end-points 0 and x.

Now c-quasiconvexity follows, if we show that

$$\theta \min\left(x\frac{f(y)}{f(x)}, y\frac{f(x)}{f(y)}\right) + 2(x-y) \leqslant c\sqrt{x^2+y^2-2xy\cos\theta}.$$

For fixed x and y, $\min\{xf(y)/f(x), yf(x)/f(y)\} \leqslant \sqrt{xy}$. Hence it suffices to show that

$$\theta^2 xy + 4\theta(x-y)\sqrt{xy} + 4(x-y)^2 + 2c^2 xy\cos\theta \leqslant c^2(x^2+y^2).$$

Since the case $y=0$ is clear, we set $s := x/y \geqslant 1$ and divide through by xy, obtaining:

$$\theta^2 + 4(\sqrt{s}-\sqrt{1/s})\theta + 4(\sqrt{s}-\sqrt{1/s})^2 + 2c^2\cos\theta - c^2(s+1/s) \leqslant 0.$$

The derivative of the left-hand side with respect to s is positive when

$$2\theta(s+1) \geqslant \left(c^2-4\right)\sqrt{s}(s-1/s)$$

or, equivalently, when $\sqrt{s}-\sqrt{1/s} \leqslant 8\theta/\pi^2$. Hence the only zero of the derivative is a maximum, and we have

$$\begin{aligned}&\theta^2+4\left(\sqrt{s}-\sqrt{1/s}\right)\theta+4\left(\sqrt{s}-\sqrt{1/s}\right)^2+2c^2\cos\theta-c^2(s+1/s)\\&\quad\leqslant \left(1+16\pi^{-2}\right)^2\theta^2+2c^2\cos\theta-2c^2\left(32\theta^2/\pi^4+1\right).\end{aligned}$$

To see that the last expression in the inequality is less than zero, we use the expression $\pi^2/4+4$ for c^2:

$$\left(1+16\pi^{-2}\right)^2\theta^2+2\left(\pi^2/4+4\right)\left(\cos\theta-32\theta^2/\pi^4-1\right) \leqslant 0.$$

When we divide by $1+16\pi^{-2}$, we see that this is equivalent to $\theta^2 \leqslant \pi^2(1-\cos\theta)/2$, which concludes the proof. □

Remark 4.1. The first part of the proof of the previous theorem shows that for the universal constant c for which every ρ_M with $M(x,y)=f(x)f(y)$ is c-quasiconvex is at least 2. For if x and y are on the same ray emanating from the origin then clearly the segment of the ray between x and y is the geodesic. Moreover the above derivation up to

$$\int_y^x \frac{dz}{f(z)^2} \leqslant \frac{2(x-y)}{f(x)f(y)}$$

is sharp. Hence $c \geqslant 2$, as claimed.

Metrics that are 1-quasiconvex are particularly interesting, since in these metric spaces any two points can be connected with a path γ with $\ell_M(\gamma)=\rho_M(x,y)$. The next lemma shows that, except for the Euclidean distance, there are no 1-quasiconvex M-relative metrics in $\mathbb{R}^n$.

Proposition 4.1. *Let M be moderately increasing. Then ρ_M is a 1-quasiconvex metric in $\mathbb{R}^n$ if and only if $M \equiv c > 0$.*

Proof. In this proof we will write r for re_1, etc. If $M \equiv c > 0$ then clearly ρ_M is 1-quasiconvex. Assume conversely that ρ_M is 1-quasiconvex. Consider the 1-quasiconvex path γ, connecting $-r$ and r, where $r>0$.

We assume without loss of generality that γ crosses the positive e_2-axis. Let $b \in \mathbb{R}^+$ be such that γ crosses the e_2-axis in be_2. Then, by the triangle (in)equality,

$$\frac{2r}{M(r,r)}=\rho_M(-r,r)=\rho_M(-r,be_2)+\rho_M(be_2,r)=\frac{2\sqrt{r^2+b^2}}{M(r,b)}$$

or, equivalently, $M(r,b) = \sqrt{1+(b/r)^2}M(r,r)$. Suppose that $b \neq 0$. Then $M(r,b) > M(r,r)$ and hence $b > r$ since M is increasing. Therefore $(b/r)M(r,r) \geqslant M(r,b)$ since M is moderately increasing. It the follows that

$$\frac{b}{r}M(r,r) \geqslant M(r,b) = \sqrt{1+(b/r)^2}M(r,r)$$

and so $b/r \geqslant \sqrt{1+(b/r)^2}$, which is impossible, hence $b = 0$.

We thus conclude that the path connecting $-r$ and r is the segment $[-r,r]$. By considering the triangle equality for a point a, with $a < r$, on the path we find that $M(r,a) = M(r,r)$. We then consider again three distinct points y, z and x on $[0,r)$ in this order. The triangle equality becomes

$$\frac{|x-y|}{M(x,x)} = \frac{|x-z|}{M(x,x)} + \frac{|z-y|}{M(z,z)},$$

hence $M(x,x) = M(z,z)$. But then $M(x,y) = M(x,x) = M(z,z) = M(z,w)$ (assuming $x \geqslant y$ and $z \geqslant w$, similarly otherwise) and we conclude $M(x,y) = c$ for $x, y \leqslant r$. Since $r > 0$ was arbitrary, it follows that $M(x,y) \equiv c$ for all $x, y \in \mathbb{R}^+$. $\square$

Remark 4.2. Note that for $M(x,y) = xy$, ρ_M is 1-quasiconvex in $\overline{\mathbb{R}^n} \setminus \{0\}$, where $\overline{\mathbb{R}^n}$ denotes the Möbius space $\mathbb{R}^n \cup \{\infty\}$, see, e.g., [11, Chapter 1]. This was shown in the proof of Theorem 1.3.

Remark 4.3. Note that the question of when a generalized relative metric of the type introduced in [4, Section 6] is quasiconvex is not directly answered by the results in this section. However, since the quasiconvexity of either the j_G metric or Seittenranta's metric, which are both generalized relative metrics, characterize uniform domains [7, 4.3–4.5], this question is clearly of interest.

5. A generalized hyperbolic metric

In this section, we will introduce the hyperbolic metric, show how our method can be used to generalize the hyperbolic metric in one setting but not in another. We use a separate method to deal with the latter case, thus solving a problem from [11, Remark 3.29].

The hyperbolic metric can be defined in several different ways, for a fuller account the reader is referred for instance to [11, Section 2]. One possible definition of the hyperbolic metric, ρ, is

$$\rho(x,y) := 2\,\mathrm{arsh}\left(\frac{|x-y|}{\sqrt{1-|x|^2}\sqrt{1-|y|^2}}\right) \tag{4}$$

for $x, y \in B^n$. An important property of the hyperbolic metric is that it is invariant under Möbius mappings of B^n. The group formed by these Möbius mappings is denoted by $GM(B^n)$. The next lemma shows that this is essentially the only generalized relative metric (in the sense of [4, Section 6]) that is Möbius invariant.

Lemma 5.1. *Let $M(x,y)=f(x)f(y)$ with $f(0)=1$. Then ρ_M is invariant under all mappings in $GM(B^n)$ if and only if $f(x)=\sqrt{1-x^2}$.*

Remark 5.1. Note that here $f(x)$ is defined only for $x \in [0,1)$. Therefore ρ_M is not exactly an M-relative metric in the sense defined in Section 1. The interpretation is nevertheless clear.

Proof. The "if" part says essentially that the hyperbolic metric is Möbius invariant, as is seen from Eq. (4), and is hence clear, see, e.g., [11, 2.49]. Assume, conversely, that ρ_M is invariant under all mappings in $GM(B^n)$.

Fix $0<r<1$ and set $d := r\sqrt{1-r^2}$. Then $d<2r$ and we may choose points $x, y \in B^n$ with $|x|=|y|=r$ and $|x-y|=d$. Let g be a Möbius mapping in $GM(B^n)$ which maps y onto the origin. It follows from [11, 2.47], that $|g(x)|=r$. Hence by Möbius invariance,

$$\frac{d}{f(r)^2}=\frac{|x-y|}{f(|x|)^2}=\frac{|g(x)-0|}{f(|g(x)|)f(0)}=\frac{r}{f(r)}$$

and so $f(r)=d/r=\sqrt{1-r^2}$. $\square$

The classical definition of the hyperbolic metric makes sense only in the unit ball and domains Möbius equivalent to it (for $n \geqslant 3$). There are however various generalizations of the hyperbolic metric to other domains. The best known of these is probably the quasihyperbolic metric that we met in Section 4. The quasihyperbolic metric is within a factor of 2 from the hyperbolic metric in the domain B^n [11, Remark 3.3].

Seittenranta's cross ratio metric is another generalization of the hyperbolic metric, with the advantage that it equals the hyperbolic metric in B^n. The reader may recall that we showed in [4, Corollary 6.5], that Seittenranta's metric can be interpreted as $\delta_G^{-\infty}$ in the one-parameter family δ_G^p, $\operatorname{card}\partial G \geqslant 2$,

$$\delta_G^p(x,y) := \log\bigl\{1+\rho'_{M,G}(x,y)\bigr\}$$

with $M=\max\{1,2^{-1/p}\}A_p$, where A_p is the power-mean defined in Section 4 and

$$\rho'_{M,G}(x,y)=\sup_{a,b\in\partial G}\frac{1}{M(|x,y,a,b|,|x,y,b,a|)}.$$

Here the cross-ratio

$$|a,b,c,d| := \frac{q(a,c)q(b,d)}{q(a,b)q(c,d)}$$

is defined for all points $a,b,c,d \in \overline{\mathbb{R}^n}$, see also [11, Chapter 1].

Seittenranta's metric is the generalization of the logarithmic expression for the hyperbolic metric given in [11, Lemma 8.39]. We now move on to study a generalization starting from the expression based on the hyperbolic cosine [11, Lemma 3.26]:

$$\rho_G(x,y) := \operatorname{arch}\Bigl\{1+\sup_{a,b\in\partial G}|a,x,b,y||a,y,b,x|/2\Bigr\}. \tag{5}$$

This can be expressed as

$$\rho_G(x,y) := \operatorname{arch}\bigl\{1 + \bigl(\rho'_{A_0,G}(x,y)\bigr)^2/2\bigr\}, \quad \text{with } A_0(x,y) := \sqrt{xy}.$$

We note that by [4, Corollary 6.5] we know that

$$\log\bigl\{1 + \rho'_{A_0,G}(x,y)\bigr\}$$

is a metric provided $\operatorname{card}\partial G \geqslant 2$. Hence by [4, Remark 3.7] we already know that

$$\operatorname{arch}\bigl\{1 + \rho'_{A_0,G}(x,y)\bigr\}$$

is a metric when $\operatorname{card}\partial G \geqslant 2$. Hence one might speculate that the area hyperbolic cosine representation of the hyperbolic metric could be generalized to the one-parameter family

$$\rho_G^p(x,y) := \operatorname{arch}\bigl\{1 + \bigl(\rho'_{A_0,G}(x,y)\bigr)^p/p\bigr\}.$$

In what follows we will, however, restrict our attention to the case $p = 2$.

Since ρ_G has previously attracted some interest, we state some of its basic properties and give an independent proof that it is in fact a metric in most domains:

Theorem 5.1 [11, 3.25 and 3.26].

(1) *ρ_G is Möbius invariant.*
(2) *ρ_G is monotone in G, that is, if $G \subset G'$ then $\rho_{G'}(x,y) \leqslant \rho_G(x,y)$ for all $x, y \in G$.*
(3) $\rho_G(x,y) \geqslant \cosh\{(q(\partial G)q(x,y))^2\} - 1$.
(4) *For $G = B^n$ and $G = H^n$ (the upper half-space), ρ_G equals the hyperbolic metric.*

Observe that ρ_G is almost a generalized relative metric, indeed, we have

$$\rho_{\mathbb{R}^n\setminus\{0\}}(x,y) := \operatorname{arch}\left(1 + \frac{|x-y|^2}{2|x||y|}\right).$$

(Note that here $\mathbb{R}^n \setminus \{0\}$ has the boundary points 0 and ∞ in $\overline{\mathbb{R}^n}$.) This expression differs from a generalized relative metric (essentially) only by the exponent 2 of $|x-y|$. However, because of this difference the question of whether it is a metric does not lend itself to the generalized metric approach of [4, Section 6].

Theorem 5.2. *The quantity ρ_G defined in* (5) *is a metric in every domain $G \subset \overline{\mathbb{R}^n}$ with* $\operatorname{card}\partial G \geqslant 2$.

Proof. It is clear that ρ_G is symmetric in its arguments. That (x,x) are the only zeros of ρ_G is also evident. Moreover, as $\operatorname{card}\partial G \geqslant 2$, ρ_G is finite. It remains to check that it satisfies the triangle inequality.

Since the supremum in the definition (5) is over a compact set (in $\overline{\mathbb{R}^n}$) it is actually a maximum. Fix x, y and z in G. Let $a, b \in \partial G$ be points such that

$$\cosh\rho_G(x,y) = 1 + |a,x,b,y||a,y,b,x|/2.$$

Define $s(a,x,y,b) := |a,x,b,y||a,y,b,x|/2$. Now

$$\operatorname{arch}\bigl(1 + s(a,x,z,b)\bigr) \leqslant \rho_G(x,z), \qquad \operatorname{arch}\bigl(1 + s(a,z,y,b)\bigr) \leqslant \rho_G(z,y).$$

Hence it suffices to prove that

$$\operatorname{arch}\big(1+s(a,x,y,b)\big) \leqslant \operatorname{arch}\big(1+s(a,x,z,b)\big)+\operatorname{arch}\big(1+s(a,z,y,b)\big). \tag{6}$$

Since s is Möbius invariant, we may assume that $a=0$ and $b=\infty$. Denote

$$s := s(0,x,z,\infty), \qquad t := s(0,z,y,\infty), \qquad u := s(0,x,y,\infty).$$

It follows that

$$s=\frac{|x-z|^2}{2|x||z|}, \qquad t=\frac{|z-y|^2}{2|z||y|}, \qquad u=\frac{|x-y|^2}{2|x||y|}. \tag{7}$$

We will first show that we may make certain assumptions regarding z by showing that if z does not satisfy these assumptions there exists another point, say z', with corresponding s' and t', such that

$$\operatorname{arch}(1+s')+\operatorname{arch}(1+t') \leqslant \operatorname{arch}(1+s)+\operatorname{arch}(1+t).$$

Hence it will clearly suffice to prove the inequality for z'.

For fixed x and y it is clear that we can move the point z so that both s and t get smaller if $|z| \leqslant \min\{|x|,|y|\}$ (since $s=(|x|/|z|)+(|z|/|x|)-2\cos\theta$ is decreasing in $|z|$ for $|z| \leqslant |x|$, and similarly for t and y). Hence we may assume that $|z| \geqslant \min\{|x|,|y|\}$. Similarly, if $|z| > \max\{|x|,|y|\}$, we can decrease s, t for fixed x and y, hence we may also assume that $|z| \leqslant \max\{|x|,|y|\}$. If $\widehat{x0y} < \pi$ we may also assume that z lies within this angle, since otherwise we may apply the transformations shown in Fig. 1 (keeping x, y and $|z|$ fixed and rotating or mirroring z according to where it started).

Since cosh is increasing, we apply it to both sides of (6) and use

$$\cosh(a+b)=\cosh(a)\cosh(b)+\sinh(a)\sinh(b)$$

to conclude that (6) is equivalent to

$$u \leqslant s+t+st+\sqrt{s^2+2s}\sqrt{t^2+2t}. \tag{8}$$

Getting rid of the square root, this equation is implied by

$$s^2+t^2+u^2 \leqslant 2(st+su+tu+stu)$$

which is equivalent to

$$(u-s-t)^2 \leqslant (4+2u)st. \tag{9}$$

Let us assume without loss of generality that $z=1$. Assume, for the time being, that 0, x, y and 1 are co-linear and that $|x|>1>|y|>0$. Then

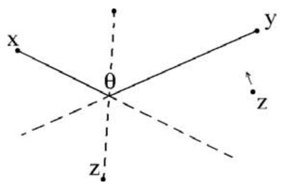

Fig. 1. The point z is between x and y.

$$s=\frac{1}{2}\left(\sqrt{|x|}-\frac{1}{\sqrt{|x|}}\right)^2, \qquad t=\frac{1}{2}\left(\sqrt{|y|}-\frac{1}{\sqrt{|y|}}\right)^2,$$
$$u=\frac{1}{2}\left(\sqrt{\frac{|y|}{|x|}}-\sqrt{\frac{|x|}{|y|}}\right)^2. \tag{10}$$

Inserting these into (9) gives

$$\left||x|+|y|+\frac{1}{|x|}+\frac{1}{|y|}-\frac{|y|}{|x|}-\frac{|x|}{|y|}-2\right|$$
$$\leqslant\left(\sqrt{\frac{|y|}{|x|}}+\sqrt{\frac{|x|}{|y|}}\right)\left(\sqrt{|x|}-\sqrt{\frac{1}{|x|}}\right)\left(\sqrt{\frac{1}{|y|}}-\sqrt{|y|}\right),$$

which is actually an equality.

Let us now consider the general case in which 0, x, y and 1 are no longer necessarily co-linear. Denote s, t and u from (10) by s_0, t_0 and u_0, respectively, and let s, t and u be as in (7). Denote

$$\delta_s:=s-s_0=(1-\cos\theta), \qquad \delta_t:=t-t_0=(1-\cos\phi),$$
$$\delta_u:=u-u_0=\bigl(1-\cos(\theta+\phi)\bigr),$$

where $\theta:=\widehat{x01}$ and $\phi:=\widehat{10y}$. Inserting $s=s_0+\delta_s$, etc., into (9) and canceling the equality $(s_0+t_0-u_0)^2=2(2+u_0)s_0t_0$ leads to

$$2(s_0+t_0-u_0)(\delta_s+\delta_t-\delta_u)+(\delta_s+\delta_t-\delta_u)^2$$
$$\leqslant 2\delta_u st+2(2+u_0)(t_0\delta_s+s_0\delta_t+\delta_s\delta_t)$$

which is equivalent to

$$(2s_0+\delta_s)(\delta_s-\delta_t-\delta_u)+(2t_0+\delta_t)(\delta_t-\delta_s-\delta_u)+(2u_0+\delta_u)(\delta_u-\delta_s-\delta_t)$$
$$\leqslant 2(stu-s_0t_0u_0). \tag{11}$$

We will first show that

$$\delta_s(\delta_s-\delta_t-\delta_u)+\delta_t(\delta_t-\delta_s-\delta_u)+\delta_u(\delta_u-\delta_s-\delta_t)\leqslant 0. \tag{12}$$

Note first that $\delta_s\geqslant 0$, $\delta_t\geqslant 0$ and $\delta_u\geqslant 0$. Now either all the parentheses are negative or $\delta_u-\delta_s-\delta_t\geqslant 0$, since $\delta_u\geqslant\delta_s,\delta_t$. In the latter case, the left-hand side of the inequality is increasing in δ_u. Since δ_s, δ_t and δ_u are squares of the sides of a triangle, we see that

$$\delta_u\leqslant\delta_s+\delta_t+2\sqrt{\delta_s\delta_t}.$$

Hence it suffices to check (12) for the maximal δ_u, in which case it is an equality.

Let us then continue from (11), using (12), rearranging and dividing by 2:

$$\delta_s(s_0-t_0-u_0)+\delta_t(t_0-s_0-u_0)+\delta_u(u_0-s_0-t_0)\leqslant stu-s_0t_0u_0.$$

Since $\delta_s,\delta_t\geqslant 0$, it follows that $stu-s_0t_0u_0\geqslant s_0t_0\delta_u$. We will then complete the proof by showing that

$$\delta_s(s_0-t_0-u_0)+\delta_t(t_0-s_0-u_0)+\delta_u(u_0-s_0-t_0-s_0t_0)\leqslant 0.$$

We may assume that (8) holds with equality for s_0, t_0 and u_0; that is

$$u_0 = s_0 + t_0 + s_0 t_0 + \sqrt{s_0^2 + 2s_0}\sqrt{t_0^2 + 2t_0}.$$

Therefore it suffices to show that

$$(\delta_u - \delta_s - \delta_t)\sqrt{s_0^2 + 2s_0}\sqrt{t_0^2 + 2t_0} \leqslant 2(t_0\delta_s + s_0\delta_t) + (\delta_s + \delta_t)s_0 t_0. \tag{13}$$

By the formula for the cosine of a sum we have, from the definition,

$$\delta_u = \delta_s + \delta_t - \delta_s\delta_t + \sqrt{(2\delta_s - \delta_s^2)(2\delta_t - \delta_t^2)} \leqslant \delta_s + \delta_t + \sqrt{(2\delta_s - \delta_s^2)(2\delta_t - \delta_t^2)}.$$

Then (13) follows if we can show that

$$\sqrt{2\delta_s - \delta_s^2}\sqrt{2\delta_t - \delta_t^2}\sqrt{s_0^2 + 2s_0}\sqrt{t_0^2 + 2t_0} \leqslant (2 + s_0)\delta_s t_0 + (2 + t_0)\delta_t s_0.$$

Let us square this equation and subtract $2\delta_s\delta_t s_0 t_0(2 + s_0)(2 + t_0)$ from both sides:

$$\bigl(2 - 2(\delta_s + \delta_t) + \delta_s\delta_t\bigr)\delta_s\delta_t(2 + s_0)(2 + t_0)s_0 t_0 \leqslant \delta_s^2 t_0^2(2 + s_0)^2 + \delta_t^2 s_0^2(2 + t_0)^2.$$

Divide both sides by $\delta_s\delta_t(2 + s_0)(2 + t_0)s_0 t_0$:

$$2 - 2(\delta_s + \delta_t) + \delta_s\delta_t \leqslant a + 1/a, \quad \text{where } a := \frac{\delta_s(2 + s_0)t_0}{\delta_t s_0(2 + t_0)}$$

(this is OK, since the cases where $\delta_t = 0$ or $s_0 = 0$ are trivial). Now $a + 1/a \geqslant 2$ so it suffices to show that $\delta_s\delta_t \leqslant 2(\delta_s + \delta_t)$ or, equivalently,

$$\frac{1}{2} \leqslant \frac{1}{\delta_s} + \frac{1}{\delta_t}.$$

But since $\delta_s, \delta_t \leqslant 2$ directly from the definition, this is clear. □

Acknowledgments

I thank Matti Vuorinen for numerous comments and suggestions as well as Glen D. Anderson and Pentti Järvi for their comments on earlier versions of this manuscript.

References

[1] A. Berglund, The p-relative distance is a metric, SIAM J. Matrix Anal. Appl. 21 (1999) 699–702.
[2] L.E. Elsgolc, Calculus of Variations, Internat. Ser. Monogr. Pure Appl. Math., vol. 19, Addison–Wesley, Reading, MA, 1961.
[3] F.W. Gehring, O. Martio, Lipschitz classes and quasiconformal mappings, Ann. Acad. Sci. Fenn. Ser. A I Math. 10 (1985) 203–219.
[4] P.A. Hästö, A new weighted metric: the relative metric I, J. Math. Anal. Appl. 274 (2002) 38–58.
[5] P.A. Hästö, A monotonicity property of ratios of symmetric homogeneous means, J. Ineq. Pure Appl. Math. 3 (5) (2002), Article 71.
[6] R.-C. Li, Relative perturbation theory. I. Eigenvalue and singular value variations, SIAM J. Matrix Anal. Appl. 19 (1998) 956–982 (electronic).
[7] P. Seittenranta, Möbius-invariant metrics, Math. Proc. Cambridge Philos. Soc. 125 (1999) 511–533.

[8] K.B. Stolarsky, Generalizations of the logarithmic mean, Math. Mag. 48 (1975) 87–92.
[9] J. Väisälä, Lectures on Quasiconformal Mappings, Springer-Verlag, Berlin, 1971.
[10] J. Väisälä, The free quasiworld. Freely quasiconformal and related maps in Banach spaces, in: Quasiconformal Geometry and Dynamics, Lublin, 1996, in: Banach Center Publ., vol. 48, Polish Acad. Sci., Warsaw, 1999, pp. 55–118.
[11] M. Vuorinen, Conformal Geometry and Quasiregular Mappings, Lecture Notes in Math., vol. 1319, Springer-Verlag, Berlin, 1988.